# Koppers, Mary Margaret

| | |
|---|---|
| **From:** | Ditthavong, Khouane |
| **Sent:** | Thursday, April 02, 1998 11:03 AM |
| **To:** | Koppers, Mary Margaret |
| **Subject:** | FW: NEW IQE RESPONSE |

I just received this message from Marion concerning Bill's response to ASTM's IQE. I assume it is related to the voting that you were talking about in the SCC meeting. I called Dale to ask to take a look at, but if you have additional insights, please talk to Dale. Thanks.

Khouane

---

| | |
|---|---|
| **From:** | MARION THOMPSON(SMTP:THOMPSON.MARION@epamail.epa.gov) |
| **Sent:** | Thursday, April 02, 1998 8:55 AM |
| **To:** | DRUSHNEC@DYNIET.COM; KDITTHA@DYNIET.COM |
| **Subject:** | NEW IQE RESPONSE |

WordPerfect 6.0

<WP Attachment Enclosed>Please get to dale now. Thanks. Trinky

William Telliard
(EPA or home address?)
Draft: March 31, 1998


Ms. Nancy E. Grams, Technical Contact
Interlaboratory Quantitation Estimate (IQE)
Ballot: D19.02(98-01)
ASTM - Standards Coordination
100 Barr Harbor Drive
West Conshocken, PA  19428-3959

Dear Ms. Grams:

I vote negative on the IQE.  My primary reason for voting negative is that the IQE is arbitrary as a quantitation limit while the descriptive aspects of the IQE are more appropriately considered as part of an existing ASTM standard.

The IQE is based on the selection of the lowest true value concentration estimated to equal some arbitrarily selected number times a standard deviation of replicate measurements at that concentration. While there are practical considerations that prevent the selection of some more specific multiplier, this is even more arbitrary than selection of some set number like 10. The %RSD used as a standard in the IQE is only a summary statistic.  A quantitation limit should be based on demonstrably practical criteria that relate to measurement.  Such practical criteria could include recognizable mass spectra or acceptable calibration points. The IQE does not include any demonstrably practical criteria related to measurement.

By allowing the %RSD associated with the IQE to be selected by the analysts, the IQE has demonstrated the fundamentally descriptive nature of %RSD. This proposed IQE is, in fact, merely the basis for some addition to the ASTM requirements set forth in the Standard Practice for Determination of Precision and Bias of Applicable Test Methods of Committee D-19 on Water (D2777). Concepts that could legitimately be included in D2777 include the estimation of between lab variability, modeling the relationship between measurement variability and true concentration, and experimental design issues related to modeling.

Additional comments and reasons for voting negative, associated with outline reference numbers from the draft standard, are as follows:

Section 1.3.2:  Carry-over error is generally an avoidable error source in routine analysis and should not be included as an allowable source of variability.

Section 1.5: If (1) statistically significant relationships are demonstrated between raw instrument responses and true value concentrations, (2) mass spectra are recognizable (where applicable), and (3) repeated measurements on a particular spike concentration indicate that measurement responses will not be censored when using that spike concentration as a calibration point then measurements equal to that spike concentration are quantitative, regardless of RSD.

Section 3.2.2: It's not clear what distinction is being made between different censored samples.

Section 4.2: There is no reason to believe that any of the three models presented in this section will fit a given set of real data. Considering the high concentrations targeted by this procedure, there is every reason to think the that the constant model will routinely fail to fit in a spectacular fashion; the constant model is generally applicable only in a region between zero concentration of the target analyte and somewhere slightly above the detection level. We have never observed the straight-line model presented here to fit any dataset. We are not aware of any publicly available dataset or report that documents the fit of the Rocke-Lorenzato model to the kind of between lab data required in this draft of the IQE. It is not appropriate to force analysts to explicitly consider these models as part of an ASTM standard procedure. Without considerably more research with real data than has thus far been presented, any attempt to constrain the statistical analysis of data collected is likely to cause unnecessary work and to result in demonstrably incorrect results. The services of a trained statistician will be required to develop appropriate models for each dataset, assuming that appropriate models can be found. ASTM should not establish standards that use models and procedures that have never been publicly demonstrated to work in an appropriate fashion.

Section 5.3: It's seems disingenuous to state that "The intent of the IQE is not to set reporting limits." The author of this method has participated in meetings where the group he was representing (now called the Inter-Industry Analytical Group (IIAG)) defined "quantitation" in their own terms and then applied their definition to EPA policy statements regarding the setting of effluent limitations at quantitative levels. It seems clear that the intended purpose of the IQE, through use of the Technology Transfer Act, is to set EPA reporting limits and to set them at higher levels than currently exist.

Section 6.2.1.2: I don't understand why a procedure that doesn't use raw instrument responses is constantly referring to recovery. It appears to me that the IQE is attempting to model measurement results and not recovery.

Section 6.2.1.3: It is unclear as to what constitutes a "single model" as opposed to saying that a model must be developed.

Section 6.2.2.1: The "semi-geometric" design will generally be preferred.

Section 6.2.3.1: It is possible that no measurement systems will produce the same measurement results with instrument thresholds turned off as with the thresholds turned on, even when measuring concentrations significantly greater than a quantitation limit. If the variability of a measurement system is different with the thresholds turned off than with the thresholds turned

on, then a quantitation limit determined with the thresholds turned off does not represent the quantitation capability of a measurement system when the thresholds are turned on. Four possible solutions are: (1) leave thresholds on for determination of the IQE, (2) determine the IQE with thresholds turned off and require all subsequent measurements to be made with the thresholds turned off, (3) develop a well-researched relationship between the variability in measurement systems with the thresholds turned off and with the thresholds turned on, or (4) require laboratories to purchase signal processing hardware and software for which it is known that changing thresholds does not affect variability characteristics. If the time and resources are not available to investigate solution (3), solution (1) is preferred. Further, if ASTM chooses to run these studies blind, then instructing the analysts to disable data-censoring limits is going to clearly indicate to analysts that a measurement study in being conducted.

6.2.4 It is rare for laboratories to operate blind. For example, laboratories checking the quality of products produced in manufacturing and laboratories establishing the quality of the effluent discharged from a facility will have specific knowledge of what these samples are likely to contain. It is therefore not realistic, or appropriate, to require blind samples for purposes of establishing quantitation limits.

Section 6.3.3.1: The reference to Johnson and Kotz should be (5), not (6).

Section 6.4: This section should explicitly state that the IQE determination is to be repeated if the estimated IQE is not within the range of concentrations used in the study.

8.1.1 At least one reviewer must be qualified in applied statistics and at least one reviewer must be qualified in analytical chemistry.

9.3.1 This section juxtaposes discussions without linking them, thus becoming misleading. Analyte recovery is not perfect but, even if analyte recovery were perfect, there is no reason to believe that measured values would exactly equal true values. After all, we generally measure properties of the analyte instead of mass weighing the analyte. However, even if we mass weighed a perfectly recovered sample of the analyte there would still be variation associated with the weighing process. Further, recovery conditions are part of the raw instrument responses that are converted to measurement values by use of an estimated calibration relationship. As long as these raw instrument responses are calibrated with an appropriate statistical relationship, measurement values are unbiased estimates. Recovery is generally a constant percentage, which ASTM is welcome to model.

Section 9.3.3: The RSD basis of the IQE may be conceptually pure, but it is by no means statistically rigorous. It is a totally arbitrary creation.

Section 10: As with the previous ballot, the IQE has not been demonstrated with real measurement values; the example included in the proposed IQE is based on a simulation. This is a serious deficiency. An appropriate estimate of the IQE requires that the relationship between true and measured concentrations be modeled with regard to both average response and variability. Before accepting the IQE as a standard, it should be demonstrated that appropriate

models can be fit to results from different measurement processes. A requirement for an appropriate fit is that the data must suggest or motivate the model. This requirement is not met. The models used in one of the references cited for the IQE (Gibbons, Coleman, and Maddalone 1997) do not appear to be motivated by the data to which they were applied. The addendum to this letter contains a short discussion concerning why the models used in Gibbons, Coleman, and Maddalone (1997) do not appear to be motivated by the data to which they were applied.

In summary, the IQE is arbitarary and irrelevant to quantitation. If ASTM finds the summary statistics related to between lab variation to be useful, then the appropriate place to establish such standards is within D2777.

Sincerely,

William Telliard
Member #......................

Addendum

One of the numerous technical difficulties with Gibbons, Coleman, and Maddalone (1997) is that the models used in that paper do not appear to be motivated by the variability observed in the data to which the models were applied. That is, plots of the data do not come close to indicating the curve estimated by the model. This can be discerned by careful examination of plots a, b, and c in Figure 2 on page 2074. These plots contain points that indicate the observed estimate of the standard deviation at each spike concentration along with a curve indicating the modeled estimate of the standard deviation within the range of the data. In plot a, standard deviations for measurements at or below spike concentrations of 100 ppt do not appear to be functionally related to standard deviations above 100 ppt. In plots b and c, standard deviations for measurements at or below spike concentrations of 200 ppt do not appear to be functionally related to standard deviations above 200 ppt. While it is possible that the models used in Gibbons, Coleman, and Maddalone (1997) could be justified with a large number of datasets in which the functional relationship is reasonably motivated by data, I am not aware that such a body of evidence exists. If such datasets exist, they should have been used as examples in Gibbons, Coleman, and Maddalone 1997 and in the proposed IQE.